4

The importance of mechanisms for the evolution of cooperation

Pieter van den Berg
Franz J. Weissing

Proceedings of the Royal Society B **282**: 20151382 (2015)

ABSTRACT

Studies aimed at explaining the evolution of phenotypic traits have often solely focused on fitness considerations, ignoring underlying mechanisms. In recent years, there has been an increasing call for integrating mechanistic perspectives in evolutionary considerations, but it is not clear whether and how mechanisms affect the course and outcome of evolution. To study this, we compare four mechanistic implementations of two well-studied models for the evolution of cooperation, the Iterated Prisoner's Dilemma (IPD) game and the Iterated Snowdrift (ISD) game. Behavioural strategies are either implemented by a 1:1 genotype-phenotype mapping or by a simple neural network. Moreover, we consider two different scenarios for the effect of mutations. The same set of strategies is feasible in all four implementations, but the probability that a given strategy arises due to mutation is largely dependent on the behavioural and genetic architecture. Our individual-based simulations show that this has major implications for the evolutionary outcome. In the ISD, different evolutionarily stable strategies are predominant in the four implementations, while in the IPD each implementation creates a characteristic dynamical pattern. As a consequence, the evolved average level of cooperation is also strongly dependent on the underlying mechanism. We argue that our findings are of general relevance for the evolution of social behaviour, pleading for the integration of a mechanistic perspective in models of social evolution.

INTRODUCTION

There is a long tradition in biology of separating proximate and ultimate perspectives when explaining phenotypic variation (Mayr, 1961; Tinbergen, 1963). The proximate perspective is concerned with the mechanisms that directly cause the phenotype (such as neurological and physiological processes), whereas the ultimate perspective is concerned with the emergence of the phenotype through (adaptive) evolution. In concordance with this traditional separation, knowledge about the specific mechanisms underlying phenotypes has long been regarded as inconsequential to the question of how phenotypes are shaped by evolution. Accordingly, evolutionary biologists have a strong focus on the fitness consequences of phenotypic traits, thereby largely disregarding the underlying mechanisms. Conceptualisation of evolution is often based on the implicit assumptions that genes interact in a simple way and that there is a one-to-one relationship between genotypes and phenotypes. These assumptions are convenient, since they allow a view of selection as a process directly acting on the genes in the "gene pool" of a population. Although this view has already been criticized as "beanbag genetics" more than fifty years ago (Mayr, 1959), theoretical approaches to explaining the evolution of phenotypes with an explicit focus on mechanisms are not very prominent even today.

Verbal discussions of the importance of underlying mechanisms for the dynamics and outcomes of evolutionary processes started to emerge in the literature in the 1980s (Maynard Smith *et al.*, 1985). In particular, the influential book of John Maynard Smith and Eörs Szathmáry on the "Major Transitions in Evolution" (Maynard Smith & Szathmáry, 1995) clearly showed how crucial genetic and phenotypic architecture are for the course of evolution. This view is now firmly established in the field of evo-devo (Arthur, 2002; Müller, 207), where the interplay between (developmental) mechanisms and evolution is at centre stage. Similarly, studies on gene-regulatory networks (Lozada-Chavez *et al.*, 2006; Aldana *et al.*, 2007; Cilibarti *et al.*, 2007) have revealed that network topology strongly affects both the robustness and evolvability of living systems, while recent 'integrative' models (Pfennig & Ehrenreich, 2014; Schneider *et al.*, 2014; Botero *et al.*, 2015) reveal that the mechanisms underlying phenotypic responses can be important for a full understanding of eco-evolutionary processes.

In line with these general developments, there are now strong pleas (McNamara & Houston, 2009; Fawcett *et al.*, 2013; 2014) to apply "mechanistic thinking" in evolutionary studies of animal and human behaviour as well. Yet, with some notable exceptions (Enquist & Arak, 1994; McNamara *et al.*, 1999; Taylor & Day, 2004; Gross *et al.*, 2008; Akcay *et al.*, 2009; McNally *et al.*, 2012), models for the evolution of behaviour still tend to make the "least constraining" assumptions on the genetic basis and the physiological and psychological processes underlying behaviour. When the direction and intensity of selection do not change in time and when there is a single optimal behaviour, this may not be problematic. In such a case, one would expect evolution to proceed towards the single optimum, regardless of underlying mechanisms. However, whenever there

are multiple equilibria, the situation is no longer so straightforward. And even in relatively simple social contexts, the existence of multiple equilibria is the rule rather than the exception (Van Damme, 1991; Nowak & Sigmund, 1993a; Van Doorn *et al.*, 2003). In other words, the question is not that much "Which strategy is favoured by natural selection" but rather "Which equilibrium will be achieved in the course of evolution" (Harsanyi & Selten, 1988; Samuelson, 1998; Gintis, 2000). It is conceivable that, in such a context, the mechanisms underlying behaviour may be of evolutionary importance, because mechanisms can affect the probabilities with which phenotypes arise and, hence, the likelihood of alternative evolutionary trajectories.

Here we study the evolution of behavioural strategies in two types of social interaction without clearly delineated optimal behaviour. Our question is whether, and to what extent, the mechanistic implementation of the available strategies affects the course and outcome of evolution. We consider two prototype models for the evolution of cooperation: the Iterated Prisoner's Dilemma game (IPD) and the Iterated Snowdrift game (ISD), which have been the subject of hundreds of earlier studies (IPD [Axelrod & Hamilton, 1981; Boyd & Lorberbaum, 1987; Binmore & Samuelson, 1992; Nowak & Sigmund, 1993b; Nowak & Sigmund, 1998; Doebeli & Hauert, 2005; McNamara *et al.*, 2008], ISD [Maynard Smith & Price, 1973; Sugden, 1986; Santos & Pacheco, 2005; Doebeli & Hauert, 2005; McNamara *et al.*, 2008]). In both games, the players have to decide (repeatedly) on whether to cooperate or to defect. For both players, mutual cooperation is more profitable than mutual defection. However, mutual cooperation is not easy to achieve, since defection yields a higher payoff than cooperation if the other player cooperates. The games differ in their assumption on whether defect (IPD) or cooperate (ISD) yields a higher payoff against a defector. Following the traditions of evolutionary game theory (Maynard Smith, 1982; Weibull, 1995; Hofbauer & Sigmund, 1998; Dugatkin, 1998; Gintis, 2000; Hofbauer & Sigmund, 2003; Nowak *et al.*, 2004; Imhof *et al.*, 2005; Nowak, 2006), but see Weissing (1996)], studies of the evolution of strategies in these games have overwhelmingly assumed a one-to-one relationship between genotypes and strategies. Here we contrast such a one-to-one implementation with a different implementation where selection does not directly act on strategies, but on the architecture (a simple neural network) underlying these strategies. In addition, we consider two genetic mechanisms that determine the probabilities with which the mutation of each strategy yields any other strategy. We will show that the evolutionary dynamics is strongly affected by both the genetic and the behavioural architecture and discuss how the different outcomes can be explained on the basis of the mutational distributions arising from the interplay between genetics and behavioural mechanisms.

THE MODEL

Games and strategies

Throughout, we will consider a Prisoners' Dilemma game and a Snowdrift game with the following payoff matrices:

$$\text{PD:}\begin{pmatrix} 3 & 0 \\ 5 & 1 \end{pmatrix}; \quad \text{SD:}\begin{pmatrix} 3 & 1 \\ 5 & 0 \end{pmatrix}.$$

The top and bottom rows give the payoffs of cooperation and defection, respectively, both for when the opponent cooperates (first column) and defects (second column). In the PD, defection always yields a higher payoff than cooperation, regardless of the action of the opponent. In the one-shot version of this game, mutual defection is therefore the only evolutionarily stable strategy (ESS). In the SD, the highest payoff is always attained by choosing the opposite action than the opponent. In this case, the one-shot game has an ESS that is characterized by a mixture of cooperation and defection.

We consider iterated versions of both games, for which the determination of all ESSs is much less straightforward than for their one-shot counterparts (see Appendix for a game-theoretical analysis). In our simulations, agents repeatedly interact for an indefinite period of rounds; after each round, the game is terminated with probability 1-*m*. The full strategy space of the iterated game is infinite-dimensional (Boyd & Lorberbaum, 1987). Here we confine the strategy space by only allowing individuals to condition their behaviour on the outcome of the previous interaction round. Since there are four possible interaction outcomes (mutual cooperation, mutual defection, and both combinations of cooperation and defection), and a strategy always prescribes one of two possible actions for each outcome (cooperation or defection), there are in total 2^4 =16 possible strategies (see Table 4.1 for a complete list). We assume that individuals are not perfect; they make both perception errors (with probability ε_p they misinterpret the behaviour of their opponent as the opposite behaviour) and implementation errors (with probability ε_i they perform the opposite behaviour than dictated by their strategy).

Behavioural and genetic architecture

Figure 4.1 shows a schematic representation of the behavioural and genetic architectures considered in this study. We consider a "1:1" behavioural architecture and an artificial neural network (ANN) architecture, which can both realize the 16 possible strategies presented in Table 4.1. The 1:1 architecture is the simplest possible architecture, in which behaviour for each of the four possible outcomes of the previous round is under the direct control of a single gene locus. Each of these four loci can only have two values: 1 (for cooperation) or 0 (for defection). In addition, a separate locus determines an individual's behaviour in the first round; this locus can take on any value of the unit interval, which correspond to the probability of cooperation in the first round.

Table 4.1. An overview of all possible pure strategies that condition their behaviour on the previous interaction. The four columns on the left show whether the strategy cooperates (1) or defects (0), for each of the four possible outcomes of the previous interaction (from left to right: mutual cooperation, having cooperated while the opponent defected, having defected while the opponent cooperated, and mutual defection). The second column from the right shows the name of the strategies that is used to refer to them in the main text. The two middle columns show the percentage of the genotype space that is associated with each strategy for the two different behavioural architectures. These percentages were obtained by generating a large number of genotypes (in the same way as generating a genotype through 'entire-genome mutation'), and subsequently determining the strategy induced by each genotype (as explained in the Appendix).

behaviour				percentage of genotype space			
CC	*CD*	*DC*	*DD*	*1:1 mapping*	*neural network*	*strategy name*	*strategy description*
0	0	0	0	6.25	40.35	*ALLD*	always defects
0	0	0	1	6.25	1.79	*desperate*	only cooperates after mutual defection
0	0	1	0	6.25	1.75	*Acon-D*	anti-conventional, shifts after playing opposite of opponent, otherwise defects
0	0	1	1	6.25	1.65	*inconsistent*	plays opposite of previous move
0	1	0	0	6.25	1.75	*con-D*	conventional, stays after playing the opposite of opponent, otherwise defects
0	1	0	1	6.25	1.65	*ATFT*	anti-tit for tat, plays opposite of opponent's previous move
0	1	1	0	6.25	0.08	*APavlov*	win, shift; lose, stay
0	1	1	1	6.25	0.98	*hopeless*	only defects after mutual cooperation
1	0	0	0	6.25	0.98	*grim*	only cooperates after mutual cooperation
1	0	0	1	6.25	0.08	*Pavlov*	win, stay; lose, shift
1	0	1	0	6.25	1.65	*TFT*	tit for tat, copies opponent's last move
1	0	1	1	6.25	1.75	*MNG*	Mr Nice Guy, only defects after ´being cheated´ (playing C while other plays D)
1	1	0	0	6.25	1.65	*consistent*	repeats its own previous move
1	1	0	1	6.25	1.75	*con-C*	conventional, stays after playing the opposite of opponent, otherwise cooperates
1	1	1	0	6.25	1.79	*willing*	only defects after mutual defection
1	1	1	1	6.25	40.35	*ALLC*	always cooperates

In the neural network architecture (Enquist & Ghirlanda, 2005; see Figure 4.1 for a graphical representation; see Appendix for a more detailed explanation), behaviour is determined through a very simple underlying structure that translates an input (the behaviours of 'self' and 'partner' in the previous interaction round) into an output (cooperation or defection). There are two input nodes, one of which receives the previous own behaviour (0 for defection, 1 for cooperation), and the other receives the previous behaviour of the opponent. The input from both these nodes is fed into two 'hidden layer' nodes, multiplied by the weights of the connections between the nodes. Each hidden layer node has a threshold; if the summed input into a hidden layer node exceeds its threshold, its output equals 1, otherwise the output is 0. Both hidden layer nodes are

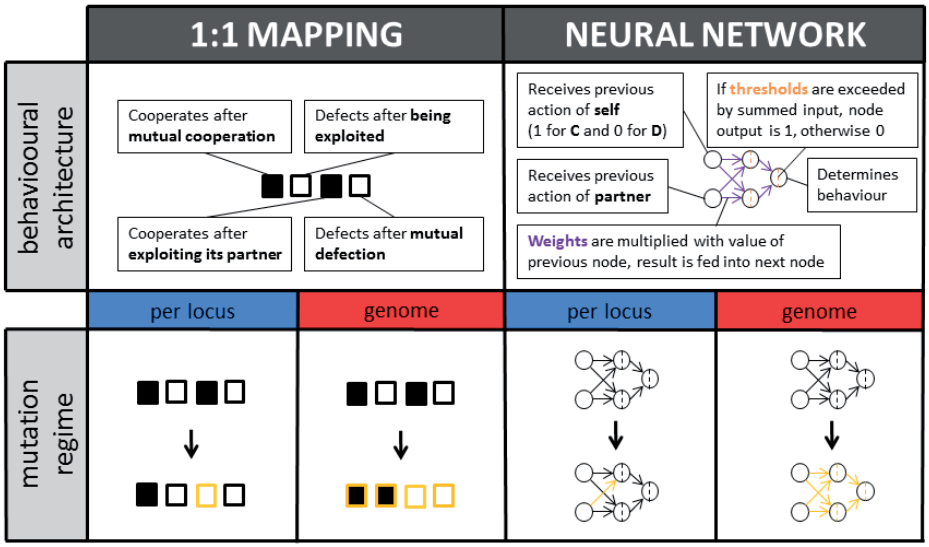

Figure 4.1. A schematic representation of the four implementations of the 16 strategies considered in this study. The top row shows illustrations of the two behavioural architectures. In the 1:1 architecture, individuals have four gene loci that each determine the behaviour (cooperate or defect) in a given round for one of the four possible outcomes of the previous round. These four loci are represented by boxes (in the example shown, black boxes represent cooperation and white boxes represent defection). In the neural network architecture, individuals have nine loci, determining the (continuous) values of six connection weights (purple) and three thresholds (orange). The network processes the input (the behaviour of 'self' or 'partner' in the previous round) into an output (cooperate or defect). In the bottom row, the two mutation regimes are illustrated for both behavioural architectures, representing the four implementations considered in this study. Under per-locus mutation, each locus mutates independently (illustrated by single loci turning yellow after the arrow). In case of whole genome mutation, all loci mutate in the event of a mutation (illustrated by all loci turning yellow after the arrow).

connected to the output node, which also has a threshold. If the total output from the hidden nodes (multiplied with the relevant connection weights) exceeds this threshold, the individual cooperates. If not, the individual defects. This way, six connection weights and three thresholds determine the strategy implemented by the network. Accordingly, the ANN is encoded by nine gene loci (that can take on any real value): one for each connection weight and one for each threshold. In addition, a tenth locus determines an individual's behaviour in the first round (as in the 1:1 implementation).

Even under the highly simplifying assumptions on the strategy set and their underlying architectures, there are many ways to implement inheritance. For example, ploidy level and linkage patterns are of considerable importance for the genetic transmission of information on strategies in sexually reproducing organisms. To keep matters as simple as possible, we here only consider

asexual populations of haploid individuals. In order to study the effect genetic factors, we restrict attention to two different mutation regimes. In both regimes, the gene locus determining the behaviour in the first round mutates independently (with probability μ_F), the mutational step size being drawn from a normal distribution with mean 0 and standard deviation σ_F. Under 'per-locus mutation', each of the other loci (4 loci in case of the 1:1 architecture and 9 in case of the ANN architecture) has a probability μ_L of giving rise to a mutation, independently of what is happening at the other loci. Under 'entire genome mutation', a mutation event (occurring with probability μ_G) affects all these loci, that is, all these loci mutate at the same time. Under both mutation regimes, mutation is implemented as drawing a random number to replace the current value of the locus (in the 1:1 architecture, this is done by drawing 0 or 1 with equal probability; in the neural network architecture, by drawing a number from a normal distribution with mean 0 and standard deviation σ_N).

Simulation setup and parameters

We simulated a population of 1,000 asexual haploid individuals, with discrete and non-overlapping generations. At the start of each generation, pairs of two individuals are formed at random. These pairs interact repeatedly, where a new round always starts with probability $m = 0.99$ (leading to an average interaction length of 100 rounds). In any given generation, all pairs play the same number of rounds. After the last round of each repeated interaction, individuals reproduce. The probability of reproducing is directly proportional to the payoff individuals accumulate over the entire repeated interaction. Population size was kept constant. After reproduction, a new cycle starts.

At the beginning of each simulation, the loci σ_F all individuals were initialized at random: initial values for the locus that determines the behaviour in the first round were drawn from a normal distribution with mean 0.5 and standard deviation 0.1; the four binary loci in the 1:1 architecture were assigned a 0 or a 1 with equal probability; and the nine loci encoding the connection weights and thresholds in the neural network architecture were assigned values that were drawn from a normal distribution with mean 0 and standard deviation σ_N. Each simulation was run for 100,000 generations. We ran 100 replicate simulations for all four combinations of the two behavioural architectures (1:1 and neural network) and the two mutation regimes (per-locus mutation and entire genome mutation). Resulting cooperation levels and strategy frequencies were calculated by averaging over all interactions in the last generation of each simulation, and then averaging those averages over all replicates.

In all simulations reported here the perception error ε_p and the implementation error ε_i were both set to 0.01; mutation probabilities (μ_L, μ_G and μ_F) were all set to 0.001, and mutational step sizes for all continuous loci (σ_N and σ_F) were set to 0.1. In the Appendix, we consider different values of these parameters in order to check for the robustness of our results.

RESULTS

Effect of architecture on the average cooperation level

We studied the evolution of cooperation in two games (the ISD and the IPD), with four different implementations (see Figure 4.1) of behavioural strategies, reflecting two scenarios concerning the underlying behavioural architecture (1:1 versus neural network), and two scenarios concerning the mutation regime (per-locus versus entire-genome). Figure 4.2 shows that in both games the evolved cooperation level is strongly affected by both the genetic and the behavioural architecture. In fact, in both games average cooperation levels were 0.4 or lower for one implementation and 0.8 or higher for another implementation. Cooperation levels were higher for the 1:1 architecture when compared to the neural network architecture in all scenarios, but the effect of the mutation regime was different between the two games. In the IPD, per-locus mutation was associated with higher levels of cooperation than whole genome mutation, whereas the opposite was true in the ISD. To understand the causes underlying these large differences, we next zoom

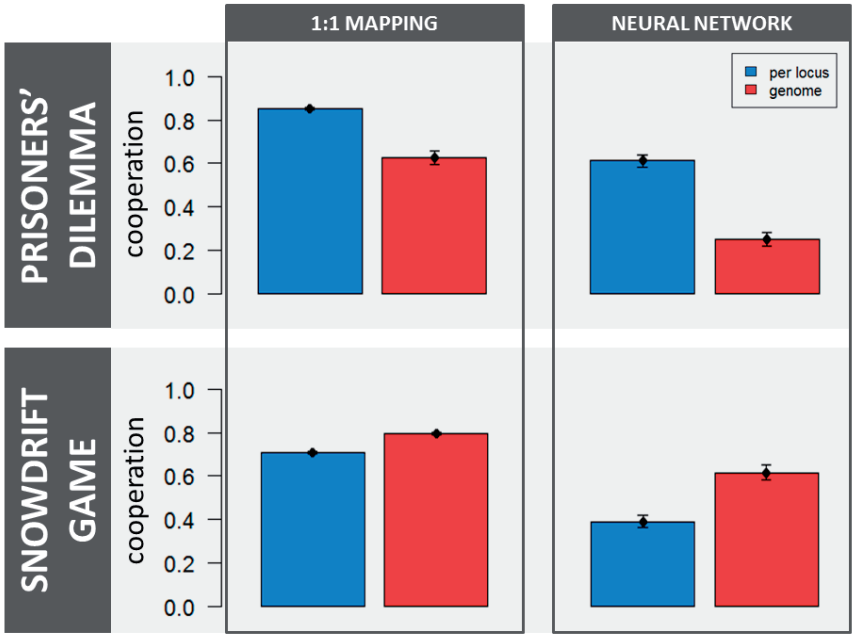

Figure 4.2. Cooperation levels in the Iterated Prisoners' Dilemma game (top) and the Iterated Snowdrift game (bottom) for all four mechanistic implementations of the 16 strategies. The bars show average cooperation levels over all interactions in the last generation, across all replicates. Error bars show standard error of the mean.

in on the evolutionary dynamics of the 16 strategies that were considered in this study (see Table 4.1 for a complete list and an explanation of strategy names).

Evolutionary dynamics in the Iterated Snowdrift game

A game-theoretical analysis of the 16 strategies in the ISD reveals that there are three evolutionarily stable strategies (see Appendix for details). ESS 1 consists of 83.3% *con-D* (conventional defector, a strategy that sticks with its previous behaviour if it played the opposite as its opponent in the previous round, and defects otherwise), together with 16.7% *ALLD*. ESS 2 consists of the pure strategy *Pavlov*. ESS 3 involves three pure strategies: 96.8% *MNG* (Mr. Nice Guy, which always cooperates, except if it cooperated while the interaction partner defected in the previous round), 2.2% *inconsistent* (which always plays the opposite to its previous move) and 1.0% *ACon-D* (unconventional defector, a strategy that changes behaviour if it played the opposite as its opponent in the previous round, and defects otherwise).

In our simulations, we recover the three ESSs above. Typically, a simulation stays at one of the ESSs for extensive periods of time, followed by a rapid shift to another ESS. In most simulations across all scenarios, ESS 1 evolved first. In some of the simulations, ESS 1 was invaded by *Pavlov*, leading to the establishment of ESS 2. In a subset of those cases, ESS 2 was ultimately invaded by *MNG*, establishing ESS 3. ESS 3 was almost never invaded, but in very rare cases could be followed by a new establishment of ESS 1. The probability of transition between two ESSs and, accordingly, the probability to find the population in any of the three ESSs strongly depends on the behavioural architecture and mutation regime (Figure 4.3): In case of the 1:1 architecture, ESS 3 is the dominant state in case of per-locus mutation, while the simulations switch between ESS 2 (attained 81% of the time) and 3 (19%) in case of whole-genome mutation. In case of a neural network architecture, the simulations either end up in ESS 1 (55%) or 3 (45%) in case of per-locus mutation and in ESS 1 (38%) and 2 (62%) in case of whole-genome mutation. In other words, the four implementations differ in their likelihood of attaining each of the ESSs, and this difference is reflected in the average cooperation levels observed in Figure 4.2 (since each ESS induces a different cooperation level; see Table 4.A1 in the Appendix).

Why do the behavioural and the genetic architecture have such a strong effect on the evolutionary outcome? This can be illustrated by considering the transition from ESS 1 to ESS 2. In ESS 1, *Pavlov* has a slight selective disadvantage when rare, but as soon as it occurs in higher frequencies it achieves a higher payoff than the strategies in ESS 1 (because the payoff of *Pavlov* against itself is high). Therefore, if *Pavlov* increases enough against the selection gradient due to mutation and genetic drift, it can invade, and ESS 2 becomes established. Clearly, the probability that *Pavlov* results from mutation of the strategies in equilibrium 1 is a crucial factor in this regard. As shown in Table 4.1 (and explained in Appendix), *Pavlov* occupies a much larger part of the genotype space (6.25%) in the 1:1 architecture than in the neural network architecture (0.08%). As a result,

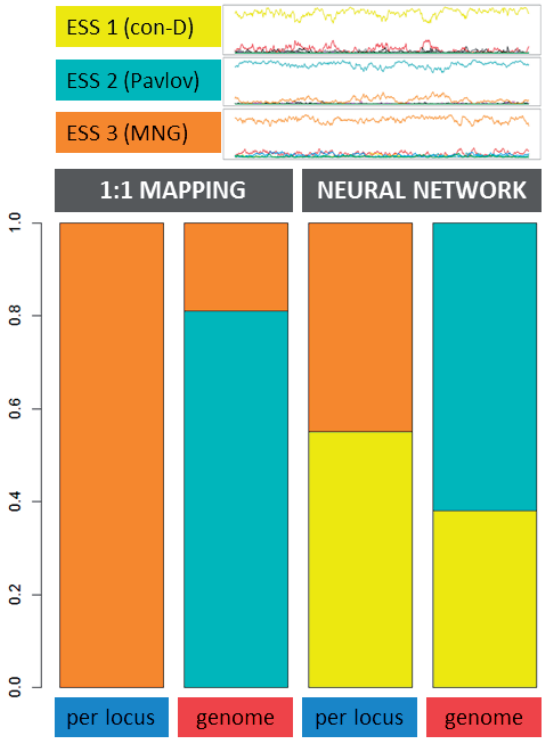

Figure 4.3. Simulation outcomes in the Iterated Snowdrift game, for all four mechanistic implementations. The three line graphs (top) show time series (2,500 generations) of typical simulation runs, each illustrating the attainment of one of the three ESSs of this game (see Table 4.A1 in the Appendix). The bar graphs (bottom) show for each scenario the fraction of 100 replicate simulations for which the last generation was in each of the three ESSs.

Pavlov almost never invades ESS 1 in the neural network architecture, whereas this does often happen in the 1:1 architecture.

Evolutionary dynamics in the Iterated Prisoners' Dilemma game

A game theoretical analysis of the 16 strategies in the IPD reveals two ESSs, both containing only a single strategy: *ALLD* and *grim* (see Appendix for details).

In our simulations, there indeed were extended periods of time in which either *ALLD* or *grim* are dominant in the population. However, in most cases, the evolution of strategies was very dynamic and often irregular. This is in line with earlier studies that also conclude that the evolutionary dynamics in an IPD is often chaotic and off-equilibrium (Rasmusen, 1989; Nowak & Sigmund, 1993a; Gosak *et al.*, 2008). Like in the ISD, both the behavioural and the genetic architecture had a strong

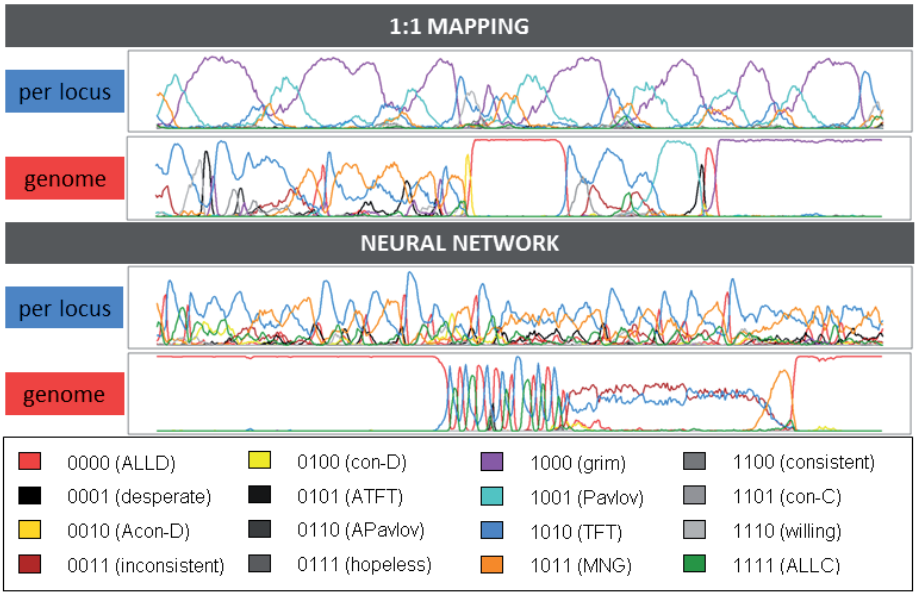

Figure 4.4. Time series of typical simulation runs in the Iterated Prisoners' Dilemma game, for all four mechanistic implementations. In each case, a time period of 2,500 generations is shown. The coloured lines represent the frequencies of the 16 different strategies.

effect on the evolutionary dynamics (see Figure 4.4). In the case of 1:1 mapping with per-locus mutation, steady cycles of *grim*, *TFT*, *MNG* and *Pavlov* were observed for all replicate simulations (this is consistent with earlier findings by Nowak & Sigmund [1993]). For whole-genome mutation, the patterns look less consistent (yet highly dynamic), including longer spells of *ALLD* domination (this explains the relatively low cooperation levels in this scenario). In the neural network architecture, per-locus mutation led to very dynamic yet fairly consistent patterns, mostly involving *TFT* and *MNG*, and infrequent *ALLD* domination spells. Entire-genome mutation typically led to long *ALLD* domination spells interspersed by short periods with both cyclical dynamics involving various strategies including *TFT*, *ALLC*, *ALLD*, *grim*, and *MNG* and non-cyclical coexistence of *TFT* and *inconsistent*.

The effect of underlying mechanisms on the evolutionary dynamics can be explained by the fact that different mechanisms induce differences in the 'mutational distance' between strategies, that is, the likelihood that a mutation in a strategy gives rise to a given alternative strategy. As an example, consider the extended periods of dominance of *grim* that were frequently observed. Those periods are typically ended by the invasion of *TFT*. *TFT* obtains a slightly worse payoff against *grim* than *grim* obtains against itself. However, *TFT* does obtain better payoffs when it happens to be paired with itself. In other words: if *TFT* can increase enough against the selection

gradient because of genetic drift and mutation, it gains a selective advantage and can invade. The probability that this occurs depends on the implementation: In the 1:1 architecture with per-locus mutation, a mutation of *grim* produces *TFT* with probability ¼. In the case of whole genome mutation, this probability is only 1/16 – this makes it considerably less likely that *TFT* obtains appreciable frequencies, and explains the extended spells of *grim* domination (Figure 4.4) and the lower degree of cooperation (Figure 4.2) in this case.

DISCUSSION

Our study demonstrates that behavioural architecture and mutation regime are of considerable importance for the dynamics and outcome of social evolution. In the IPD, we observed three types of dynamic behaviour (predictable cycles; fast and chaotic dynamics; spells of *ALLD* or *grim* domination) whose occurrence crucially depended on both behavioural architecture and mutation structure. Likewise, the prevalence of and the transitions between the three ESSs in the ISD were strongly determined by both architecture and mutation structure. In both games, the differences in evolutionary dynamics resulted in substantial differences in the average level of cooperation. These conclusions are not specific to the parameters considered in our simulations; they also hold for different payoff configurations of both games, and for a lower degree of stochasticity in the simulations (see Appendix).

The effect of mechanisms on the evolutionary dynamics was not caused by 'hard' constraints (the inability of mechanisms to produce all phenotypes), since all 16 strategies of the game were feasible in all four implementations. Yet, the mechanisms induced some 'soft' constraints on evolution, by strongly affecting the probabilities with which strategies arise by mutation (see Figure 4.5 for a schematic overview of mutation probabilities for each scenario considered in this study). Even in case of small mutation rates, the mutational distribution has a strong effect on the type of variation that can be expected to be present in a given situation. Some strategies only gain a selective advantage once they have increased beyond a certain frequency, and mutation probabilities determine the probability that this will happen. In the 1:1 architecture considered in this study, each strategy has an equal probability to result from a randomly generated genotype, whereas in the neural network architecture, some strategies (notably *ALLD* and *ALLC*) are much more likely to arise due to mutation than others (see Table 4.1). In the case of entire-genome mutation, the strategy of a mutant individual is independent of the strategy of its parent, and mutation probabilities therefore only depend on the behavioural architecture. In the case of per-locus mutation, the parental strategy partly determines the strategy of their mutant offspring.

We are not the first to point out that the genotype-phenotype mapping and the induced mutation structure are important for the course of evolution. In fact, the formal frameworks for modelling evolutionary dynamics can, to a certain extent, take these complexities into account. For

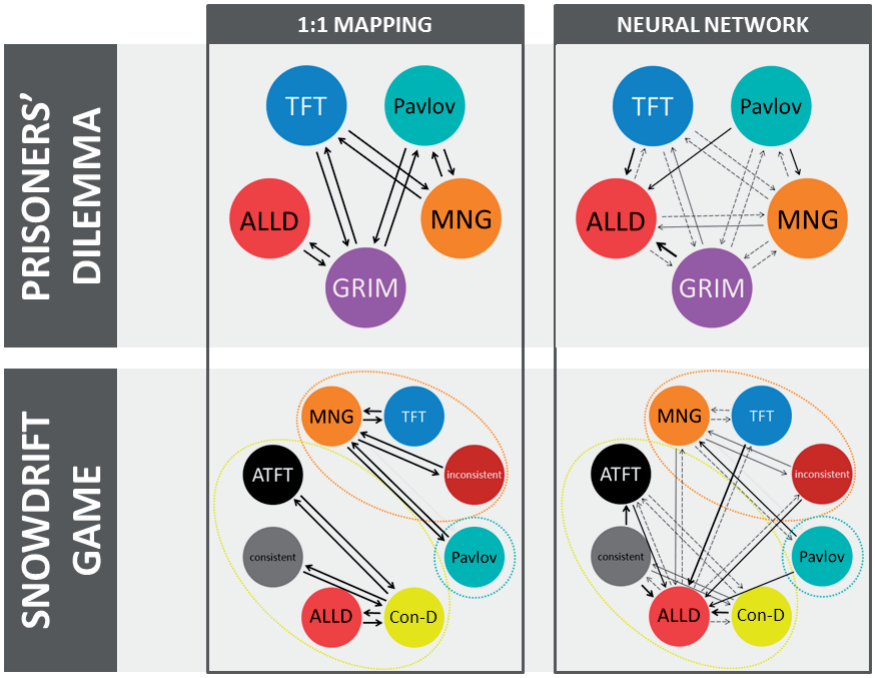

Figure 4.5. Mutational distance between the most relevant strategies in the Prisoners' Dilemma game (top) and the Iterated Snowdrift game (bottom), for both behavioural architectures and the case of per-locus mutation. An arrow pointing from one strategy to another indicates that a mutation of the former strategy has probability of larger than 0.001 to yield the latter strategy. A probability of > 0.05 is indicated by a solid arrow (the thickness of the arrow is proportional to the probability). In the 1:1 model, each strategy can mutate to four other strategies with equal probability, so the arrows in the mutation maps for the 1:1 model all represent a probability of 0.25. To calculate these probabilities, we first generated a large number of random genotypes (in the same way as generating a genotype through 'entire-genome mutation'), and determined their corresponding strategy. Then, for each strategy, we mutated all corresponding genotypes many times, and again determined the resulting strategies.

example, the 'canonical equation' of adaptive dynamics theory includes a mutational covariance matrix (Dieckmann & Law, 1996; Metz & De Kovel, 2013), which characterizes the likelihood that a combination of phenotypic traits (like a conditional strategy) arises and potentially invades the current resident strategy. Likewise, the multivariate selection equation of quantitative genetics (Lande, 1979; Lande & Arnold, 1983) can be written in a form that it includes a matrix characterizing the covariance in phenotypic traits between parents and offspring (Rice, 2004). Sean Rice has worked out formally how these covariances arise and determine the course of evolution when phenotypic traits are the outcome of (developmental) mechanisms (Rice, 2002; 2004). We are not aware of attempts to actually derive the covariance matrices of adaptive dynamics of quantitative genetics theory on the basis of a concrete mechanistic model. Instead, theoretical studies

tend to make simplifying assumptions, such as replacing the covariance matrix by the identity matrix (e.g. Van Doorn *et al.*, 2003). Already in the case of frequency independent selection and in the absence of stochasticity, assumptions like these are not unimportant, since the covariance structure largely determines which peak of a multi-peaked fitness landscape will be reached.

There are two main reasons why we think that phenotypic covariances and, hence, the mechanisms underlying the development of phenotypes, are of particular importance for social evolution. First, selection will virtually always be frequency-dependent in this case. As a consequence, the success of each strategy will strongly depend on the context, and, in particular, on the presence of selectively favoured competitors. Accordingly, a given architecture will contribute to the stability of a given equilibrium if it makes the production of selectively favoured alternatives less likely, and it will have a destabilizing effect if the opposite is the case. Second, in case of social interactions there are typically many alternative Nash equilibria and ESSs. This is already illustrated in the IPD and the ISD with highly restricted memory considered here. Relaxing the restrictions on the strategy set would lead to a rapid increase in the number of equilibria. In fact, the Folk Theorem of game theory (Van Damme, 1991; Brembs, 1996) implies quite generally that in repeated games the set of Nash equilibrium strategies is so large that virtually any 'reasonable' outcome (in case of our IPD: any outcome between 0 and 5; in case of our ISD: any outcome between 1 and 5) can be achieved as the average outcome of an equilibrium. But also non-repeated games typically have several (and often a large number of) Nash equilibria and ESSs (Selten, 1983). In all these cases, it is to be expected that the evolutionary dynamics will be affected in a similar way by mechanisms as in the present study.

We have here focussed on situations where the evolutionary game dynamics (Hofbauer & Sigmund, 2003) are relatively simple. For the payoff structure considered, the IPD and the ISG have a small number of ESSs, and these are the only attractors of the replicator equation (see Appendix). Accordingly, mechanisms will mainly affect the transition between ESSs, as described above. It is conceivable that mechanisms have an even stronger effect in the presence of limit cycles or other non-equilibrium attractors. Such attractors regularly occur in evolutionary games (such as variants of the Rock-Scissors-Paper game; see Weissing, 1991), and it has been shown that seemingly small differences in the genetic implementation of strategies can have major effects on the evolutionary outcome (Weissing & Van Boven, 2001). In the Appendix, we show that non-equilibrium attractors can also occur in the IPD (for slightly different payoff parameters), but a thorough investigation of the interplay of genetic or behavioural architecture and non-equilibrium dynamics is beyond the scope of this study.

Our results should be mainly viewed as proof of principle that mechanisms matter for the course and outcome of social evolution. It would be premature to conclude that one of the four implementations considered in our study is more 'realistic' than the others. On purpose, we kept our assumptions on architecture as simple as possible, since this allowed us to develop a sound intuitive understanding of our results (see Figure 4.5). Because of this understanding, we are

confident that our findings are of general relevance. The development of truly 'realistic' models remains a major challenge, since the actual genetic, physiological, neurological, and psychological mechanisms behind social behaviour are still largely *terra incognita* for virtually all organisms and virtually all interaction types. For this reason, it would be premature to abandon the standard 1:1 genotype-phenotype mapping assumption in favour of (for example) a neural network implementation. However, whatever the implementation chosen, researchers should be aware that it may have considerable implications for the course and outcome of evolution.

The evolution of social behaviour is often an intricate process, with many feedbacks at work, and many possible outcomes. We have shown that underlying mechanisms are of decisive importance in determining which outcome eventually emerges in evolution. Therefore, it is of importance that we focus more on mechanisms when trying to explain the evolution of social behaviour. Both empirical work focused on understanding mechanisms and theoretical work investigating their importance for the dynamics and course of evolution have a vital role to play in this regard.

Appendix

This appendix consists of three sections. First, we illustrate with an example how the neural network implements a strategy. Second, we present an ESS analysis of both games that we consider (the IPD and the ISD). Third, we present a sensitivity analysis showing that the main conclusions of our study hold for a number of alternative parameter settings.

1. IMPLEMENTATION OF STRATEGIES BY A NEURAL NETWORK

All sixteen strategies listed in Table 4.1 can be realized by our neural network architecture. To give some more insight in how a neural network configuration implements a strategy, Figure 4.A1 gives an example of one of the many possible networks implementing *Pavlov*. *Pavlov* cooperates after mutual cooperation or mutual defection, and defects otherwise. The implementation of *Pavlov* in the neural network architecture depends on a relatively precise relationship between the weights and thresholds of the network. For example, the sum of w_5 and w_6 must exceed t_3, but neither of the two should exceed t_3 alone. Also, w_1 and w_3 must together exceed t_1, and only one of those two weights must exceed t_1 alone, while the other must not. The same necessary relationships hold for weights w_2 and w_4 and threshold t_2. In addition, both t_1 and t_2 must be negative. The relatively restrictive conditions needed for a network to correspond to *Pavlov* give an intuitive understanding of why only a small percentage of the genotype space is associated with this strategy (see Table 4.1).

When considering the network in Figure 4.A1, it also becomes easier to understand why the strategies *ALLD* and *ALLC* are so likely to result from a random configuration of the network. If the value of t_3 in this network is changed from 4 to 6, then this threshold can never be exceeded, meaning this network will always defect. More generally, all networks for which $t_3 > w_5 + w_6$ and $t_3 > 0$ implement *ALLD*, regardless of the values of the other weights and thresholds. Similarly, all networks for which $t_3 < w_5 + w_6$ and $t_3 < 0$ implement *ALLC*. However, this only one of many ways that the network in Figure 4.A1 could mutate towards a network implementing *ALLD*. For example, since both t_1 and t_2 must be exceeded for this network to cooperate, any mutation that increases the value of either of these two thresholds so that it cannot be exceeded by the preceding weights would result in an *ALLD* network.

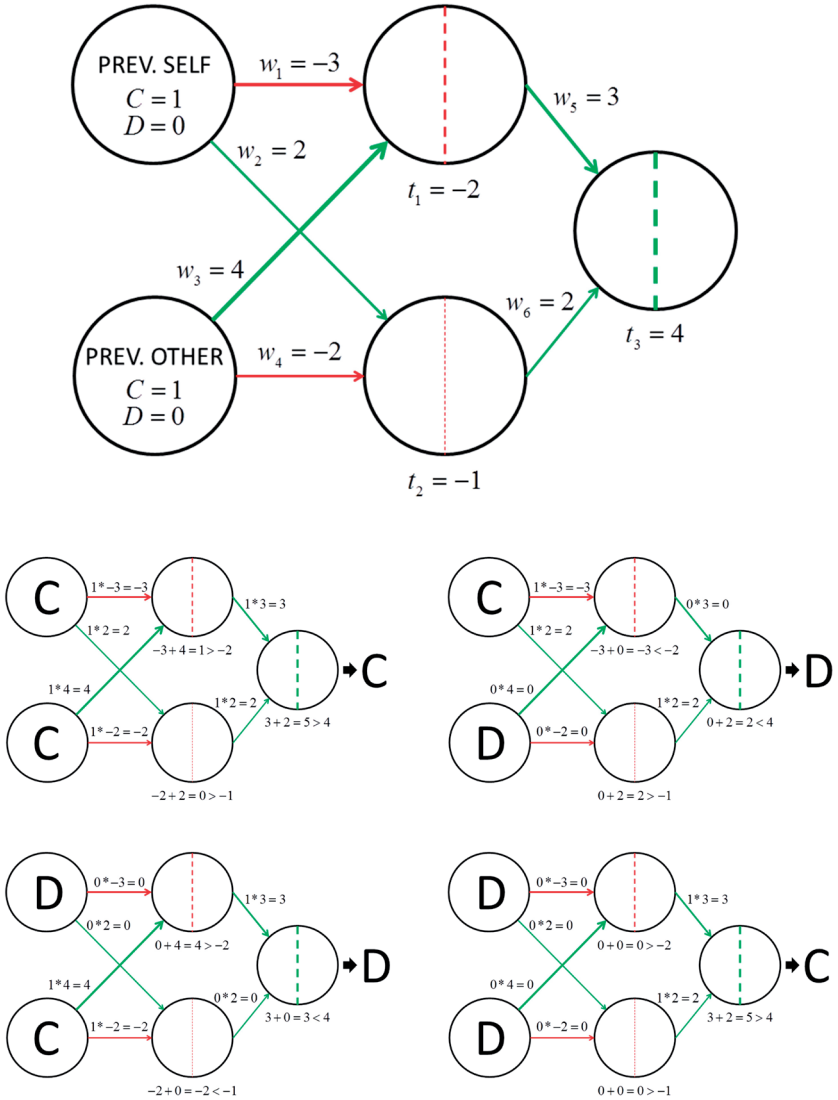

Figure 4.A1. A neural network corresponding to the *Pavlov* strategy. The network on top shows the values of all the weights (w_1 – w_6) and thresholds (t_1 – t_3); green lines depict positive values, red lines depict negative values, and their thickness indicate their absolute values. The input nodes receive the previous own decision (top) and the previous decision of the interaction partner (bottom); 1 for cooperation and 0 for defection. The four networks on the bottom show in detail how the output of the network is generated for each of the four possible outcomes of the previous round. The values of the input nodes are multiplied with weights w_1 to w_4; the resulting values are summed and fed into the hidden layer. If the summed values exceed the respective threshold values of the hidden layer, their output is 1; otherwise, it is 0. Those values are then multiplied with weights w_5 and w_6, and the sum of the resulting values is fed into the output node. If this value exceeds the threshold of the output node, the network cooperates; otherwise, it defects.

2. ESS ANALYSIS OF THE IPD AND THE ISD

Repeated games like the Iterated Prisoners' Dilemma game (IPD) or the Iterated Snowdrift Game (ISD) have a huge strategy set, and even for games as simple as these a full game theoretical analysis has not yet been achieved. Such an analysis is a formidable task, since the number of Nash equilibria is huge (Van Doorn *et al.*, 2003). In fact, the 'Folk Theorem' of game theory implies that *any* reasonable outcome can be realized by a Nash equilibrium of the iterated game (for details see Van Damme, 1991). On the other hand, no pure strategy is evolutionarily stable in the IPD or the ISD (e.g. Boyd & Lorberbaum, 1987).

Here, we consider simpler versions of the IPD and the ISD where the strategy set is restricted to 16 pure strategies with limited memory (only the moves in the previous round are memorized). These versions of the IPD and the ISD have been the subject of many studies, but to our best knowledge a full characterization of all Nash equilibria or all evolutionarily stable strategies (ESS) has never been given. This is understandable, since even under the restriction to 16 pure strategies the game theoretical analysis remains a challenge. Here we show how to classify all ESSs with support 1, 2 and 3 of the two repeated games for the payoff configuration considered in the main text. A *Mathematica* file with all calculations is available upon request. An overview of all ESSs with support 1, 2 and 3 is shown in Table 4.A1. By means of numerical iterations based on the replicator equation, we demonstrate that most likely there are no other ESSs with a larger support.

Construction of the 16x16 payoff matrix

As a first step, we determine the payoff matrix of the game by calculating the expected payoffs for any pair of pure strategies. Any such pair of strategies induces a sequence of transitions between the four different 'cooperation states' of a game round (mutual cooperation, mutual defection,

Table 4.A1. The evolutionarily stable strategies (ESS) identified by the game theoretical analysis of the IPD and the ISD. In case of the ISD, the three ESSs are numbered in line with the three equilibrium outcomes of the simulations described in the main text. The last column gives the average cooperation level in each equilibrium. A full description of all the strategies can be found in Table 4.1 of the main text.

Game	*ESS*	*Strategies*	*Fraction*	*Cooperation*
IPD	1	*ALLD*	1.000	0.010
	2	*grim*	1.000	0.013
ISD	1	*con-D*	0.833	0.180
		ALLD	0.167	
	2	*Pavlov*	1.000	0.943
	3	*MNG*	0.968	0.698
		inconsistent	0.022	
		Acon-D	0.010	

cooperation-defection, defection-cooperation). In the absence of errors, this sequence would be deterministic and mainly dependent on the combination of moves in the first round. The inclusion of perception and implementation errors, however, turns the transition between states into a stochastic process with a well-defined 4x4-matrix of transition probabilities between cooperation states. The right eigenvector corresponding to the dominant eigenvalue of this matrix (which is typically the only positive eigenvalue of the matrix) is proportional to the stationary distribution over states generated by the interaction of the two pure strategies. The four elements of the (normalized) eigenvector correspond to the fraction of time spent in each of the four cooperation states in an infinitely repeated interaction. Weighing the four elements of the payoff matrix of the one-shot game by these elements and summing up the results therefore yields the expected per-round payoff for each of the two strategies. Notice that this pair of payoffs does not depend on the initial pair of moves (these are irrelevant from a long-term perspective), but reflects the perception and implementation errors made by the players (both were kept at 0.01, as in the simulations). All subsequent ESS calculations are based on the 16x16 payoff matrix that results by applying the above recipe to all pairs of pure strategies.

Determination of all pure-strategy ESSs

It is straightforward to characterize all pure-strategy Nash equilibria: A pure strategy *i* is a Nash equilibrium if no alternative pure strategy *j* attains a higher payoff against *i* than *i* attains against itself. If *i* is a 'strong' Nash equilibrium (any alternative pure strategy *j* attains a lower payoff against *i* than *i* attains against itself), then *i* is an ESS (Maynard Smith, 1982; Selten, 1983; Hofbauer & Sigmund, 1988; Van Damme, 1991; Weissing, 1991). It turned out that the IPD has two pure-strategy Nash equilibria (*grim* and *ALLD*), which both are strong Nash equilibria and therefore an ESS. The ISD has a single pure-strategy Nash equilibrium (*Pavlov*), which again is a strong Nash equilibrium and an ESS.

Determination of all ESSs with two coexisting pure strategies

To calculate all ESSs with support two, we first determined all those pairs of pure strategies *i* and *j* that can mutually invade each other: *j* has a higher payoff against *i* than *i* has against itself; and *i* has a higher payoff against *j* than *j* has against itself. This condition of mutual invadability is equivalent to the existence of a mixed-strategy ESS of the restricted game with only these two pure strategies (Hofbauer & Sigmund, 1988; Weissing, 1991). This ESS can easily be calculated on the basis of the condition that the fitness of both pure strategies needs to be identical at the ESS (Maynard Smith, 1982). The ESS thus found is an ESS of the full game (with all 16 strategies) if all other strategies have a lower payoff at this two-strategy equilibrium than the two equilibrium strategies (Weissing & Van Boven, 2001). It turned out that the IPD has no ESS with support two, while the ISD has one such ESS: 83.3% *con-D* and 16.7% *ALLD*.

Determination of all ESSs with three coexisting pure strategies

The calculation of all ESSs with support three is more tedious. For each triplet *i*, *j*, and *k* of pure strategies we first checked whether the three strategies can coexist at equilibrium. This was done by checking whether the system of linear equations specifying fitness equality of the three strategies in the restricted three-strategy game has a positive solution. This solution specifies a candidate-ESS. In a second step, the 'internal' stability of this candidate-ESS (i.e., its evolutionary stability in the restricted three-strategy game) was checked, making use of a criterion for evolutionary stability in 3x3 games (Hofbauer & Sigmund, 1988; Weissing, 1991). In a third step, the 'external' stability of the candidate-ESS was determined by checking whether all other 13 pure strategies of the full 16-strategy game have a lower payoff at the candidate-ESS than the three strategies being part of the candidate-ESS (Weissing & Van Boven, 2001). It turned out that the IPD has no ESS with support three, while the ISD has one such ESS: 96.8% *MNG*, 2.2% *inconsistent*, and 1.0% *Acon-D*.

Numerical iterations based on the replicator equation

To check whether the equilibria identified in the game theoretical analysis are actually attainable and dynamically stable, and whether any other attractors are present in the system, we performed extensive numerical iterations using the replicator equation (Hofbauer & Sigmund, 1988; Weissing, 1991). To do this, we used the same 16x16 payoff matrices calculated for the game theoretical analysis of both games (described above). Starting from around $6*10^7$ different initial population constitutions, we iterated the replicator equation until an attractor was reached. For each iteration, the minimum frequency of each strategy was set to 0.001 so that invasion of any strategy was always in principle possible. The outcome of these iterations was congruent with the game theoretical analysis; all iterations (for both of the games) ended up in one of the ESSs in Table 4.A1, and each of the ESSs was commonly attained (depending on the initial conditions). The oscillating behaviour commonly observed in the individual-based simulations of the IPD was never observed in the (deterministic) numerical iterations. This suggests that those observations correspond to transient ('away from equilibrium') behaviour driven by the stochastic component of the individual-based simulations.

3. SENSITIVITY ANALYSIS

In the main text, we have shown that the mechanisms underlying the strategies of an evolutionary game can be of substantial importance for the evolutionary dynamics and the evolutionary outcome. For ease of representation, all simulations were conducted for one set of parameters. In this section, we show that our main conclusion also holds for a number of altered parameter

settings. Specifically, we consider versions of our model with a reduced mutation rate and with an altered payoff configuration in both games.

Reduced mutation rate

The simulations of the IPD discussed in the main text commonly exhibited highly dynamical behaviour, even though no non-equilibrium attractors were identified in the game theoretical analysis or numerical iterations of the replicator equation (see Appendix section 2). It is likely that this discrepancy is caused by the stochastic component of the simulations. The degree of stochasticity can be reduced by increasing the population size or by decreasing the mutation rate. Here, we investigate how our evolutionary simulations are affected by a reduction of the

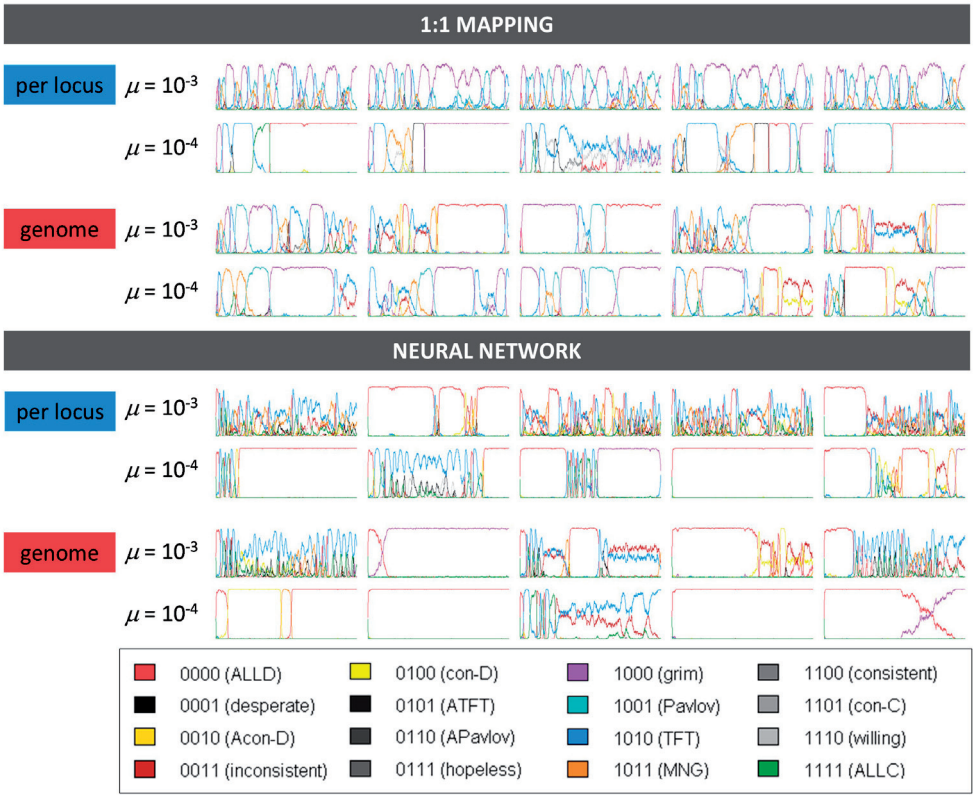

Figure 4.A2. Typical simulation runs for the IPD, for different behavioural architectures, mutation regimes, and mutation rates. For each parameter combination, excerpts of five replicate simulation runs are shown. Each excerpt comprises a time period of 2,000 generations. The excerpts were chosen from 100 replicate simulation runs (comprising 100,000 generations each) to give an impression of the 'typical' dynamics observed for each parameter combination. Graphs of all 100 replicate simulation runs for each parameter combination are available upon request.

mutation rate from 10^{-3} to 10^{-4}. In addition, we also give an idea of how replicate simulation runs can differ from each other.

Figure 4.A2 shows typical simulation runs of the IPD for each of the scenarios (behavioural architecture and mutation regime), for both mutation rates. The game theoretical analysis identified two pure-strategy ESSs for this game: *ALLD* and *grim* (see Table 4.A1). Most simulation runs consist of periods of stasis (with one or two strategies in an equilibrium-like situation), followed by periods of strong fluctuations. As a rule, the periods of stasis are longer in case of a lower mutation rate, but even in that case, highly dynamical periods are common. During the periods of stasis, one of the two ESSs (*ALLD* or *grim*) is typically the predominant strategy. However, many simulation runs included prolonged periods of stasis dominated by the non-ESS strategies *Pavlov* or *MNG*, or *con-D* or static periods where *TFT* and *inconsistent* coexisted in almost equal frequencies. Interestingly, these deviations from the game theoretical ESS predictions were mainly observed in simulations with a low mutation rate.

Besides these effects of a lowered mutation rate on the simulation dynamics, Figure 4.A2 clearly shows that lowering the mutation rate does not change our main conclusion that underlying

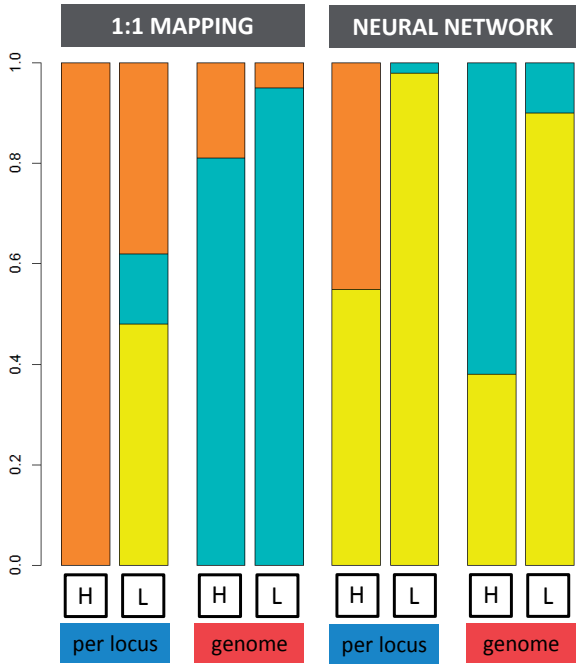

Figure 4.A3. Evolutionary outcome of the simulations of the ISD for the different mechanistic implementations and mutation rates. The left bars of each pair show high mutation rates (H, $\mu = 0.001$), bars on the right show low mutation rates (L, $\mu = 0.0001$). Bars indicate the fraction of 100 replicate simulations ending up in each of the three ESSs of the ISD (Table 4.A1) after 100,000 generations.

mechanisms strongly affect the evolutionary outcome. For example, *ALLD* domination was more commonly observed in the simulations with the neural network implementation, whereas *grim* domination was more common for the 1:1 mapping.

The game theoretical analysis of the ISD identified three ESSs (see Table 4.A1). Irrespective of the mutation rate, most simulations quickly converged to one of these ESSs; transitions between ESSs occurred, but only on rare occasions. When transitions occurred, ESS 1 was generally attained first, eventually succeeded by ESS 2, and finally followed by ESS 3 in some simulations. As expected, such transitions were less frequent in case of a lower mutation rate.

Figure 4.A3 shows how frequently each of the three ESSs was attained after 100,000 generations in 100 replicate simulations of the ISD, for each behavioural architecture, mutation regime, and mutation rate. In all cases, ESS 1 was observed more often for low mutation rate than for high mutation rate (except for the 1:1 mapping with entire genome mutation, where this ESS was never the outcome, regardless of the mutation rate). Similarly, ESS 3 was always observed more often for the higher mutation rate (except for the neural network with entire genome mutation, where this equilibrium was never observed). Besides these effects, lowering the mutation rate does not alter our main conclusion: also in the ISD the underlying mechanisms strongly affect the outcome of evolution.

Alternative payoff configurations

A full sensitivity analysis concerning the effect of the payoff parameters of the IPD or the ISD on the evolutionary outcome is beyond the scope of this study. As mentioned above, even an ESS analysis or an analysis of the attractors of the replicator equation as a function of the payoffs is quite a challenge. This is exemplified by the game theoretical analysis in section 2 of this appendix, which shows that the 16x16 payoff matrix of the two games does not have a straightforward relationship with the payoffs of the underlying one-shot game. Instead of conducting a comprehensive analysis of the effect of payoffs, we he here just give an illustration, by redoing our analysis for one alternative payoff configuration for each of the games. As will become clear, even this relatively small alteration qualitatively changes the equilibrium structure of both games, but does not affect our main conclusion that underlying mechanisms strongly affect the outcome of social evolution.

In this analysis, we replaced the payoffs for defecting while the interaction partner cooperates from 5 to 4 in both games, yielding the following payoff matrices:

$$\text{PD: } \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix}; \quad \text{SD: } \begin{pmatrix} 3 & 0 \\ 4 & 1 \end{pmatrix}.$$

First, we performed a game theoretical ESS analysis on the iterated versions of these games (as described in section 2 of this appendix); all ESSs with support 1, 2 or 3 are shown in Table 4.A2. In case of the IPD, the game theoretical analysis identifies the same pure-strategy ESSs as

in the original game (*ALLD* and *grim*), but also identifies a third (*Pavlov*). As in the original game, there are no ESSs with 2 or 3 coexisting strategies in this game. In the ISD, we also find the same pure-strategy ESS as in the original game (*Pavlov*). In addition, we find another pure-strategy ESS: *con-D*. This ESS is relatively similar to ESS 1 of the original model (that equilibrium also consisted of mostly *con-D*, but also included a relatively low frequency of *ALLD*; see Table 4.A1). These two pure-strategy ESSs are the only ESSs identified by the game theoretical analysis; ESS 3 (or a similar ESS) was not identified in this version of the ISD.

To check whether there are any other attractors in the system, we also did extensive numerical iterations of the replicator equation (see section 2 of this appendix for details). It turned out that all the ESSs described above were commonly attained; and that there is apparently no alternative stable equilibrium. Interestingly, a non-equilibrium attractor appeared in the replicator dynamics of the IPD: about 18% of all iterations converged to the stable limit cycle shown in Figure 4.A4. This cyclical attractor includes most of the 16 pure strategies. The most prominent role is for *TFT*, followed in time by *con-C* (and, in lower frequencies, *willing* and *ALLC*), then by a mix of strategies including *ATFT*, *desperate*, *con-D*, *grim*, *ALLD*, and *Pavlov*, to be followed by *TFT* again. After this second peak of *TFT*, a similar mix of strategies follows as after the first peak, although *ATFT* is absent this time, there is a small peak of *inconsistent*, and there is a longer period of a mix of *ALLD* and *grim* before the cycle starts over again. Except for this last stretch, a small frequency of *MNG* is present throughout. Only a minority of strategies do not play a role in this cyclical attractor: *Acon-D*, *APavlov*, *hopeless*, and *consistent*.

Figure 4.A5 shows typical behaviour of simulations of the IPD for the altered payoff configuration. These simulations were run for all mechanistic implementations, and also for high (10^{-3}) and low (10^{-4}) mutation rates. It is apparent that in comparison to the original game the periods of stasis are more pronounced, while the periods of fluctuation tend to be much shorter. For example, the 1:1 mapping with per-locus mutation now typically converges to the ESS *Pavlov*, while strong fluctuations were the rule in the original game (Figure 4.A2). The reduced tendency

Table 4.A2. ESSs identified by the game theoretical analysis of the IPD and ISD for the alternative payoff configuration. The last column gives the average cooperation level in each equilibrium. Note that there is also a fourth cyclical attractor in the IPD, of which the dynamics are illustrated by Figure 4.A4 A full description of all strategies can be found in Table 4.1 of the main text.

Game	*ESS*	*Strategies*	*Fraction*	*Cooperation*
IPD	1	*ALLD*	1.000	0.010
	2	*grim*	1.000	0.013
	3	*Pavlov*	1.000	0.943
ISD	1	*con-D*	1.000	0.206
	2	*Pavlov*	1.000	0.943

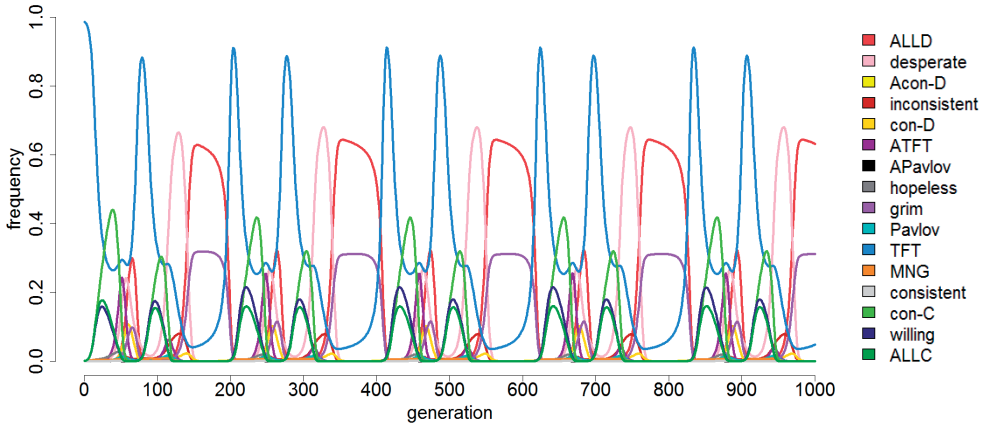

Figure 4.A4. A cyclical attractor in the IPD with alternative payoff matrix. Note that some of the colours associated with the strategies are different than in the other graphs.

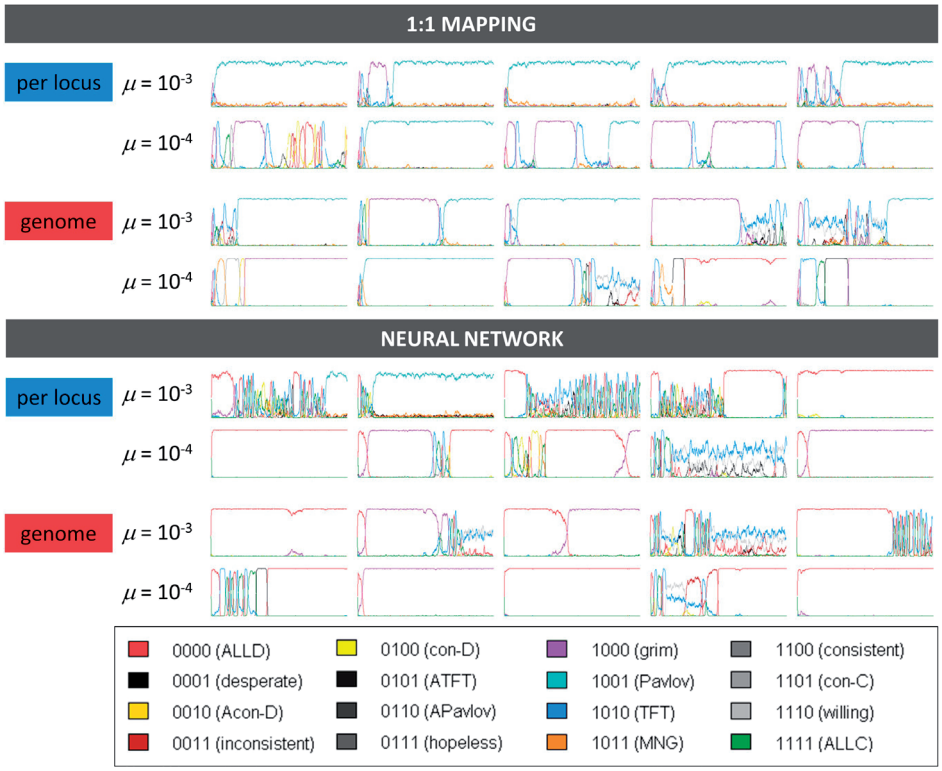

Figure 4.A5. Typical simulation runs for the IPD with altered payoff matrix, for different behavioural architectures, mutation regimes, and mutation rates. As in Figure 4.A2, representative excerpts of 100 replicate simulation runs are shown for each parameter combination. Graphs of all 100 replicate simulation runs for each parameter combination are available upon request.

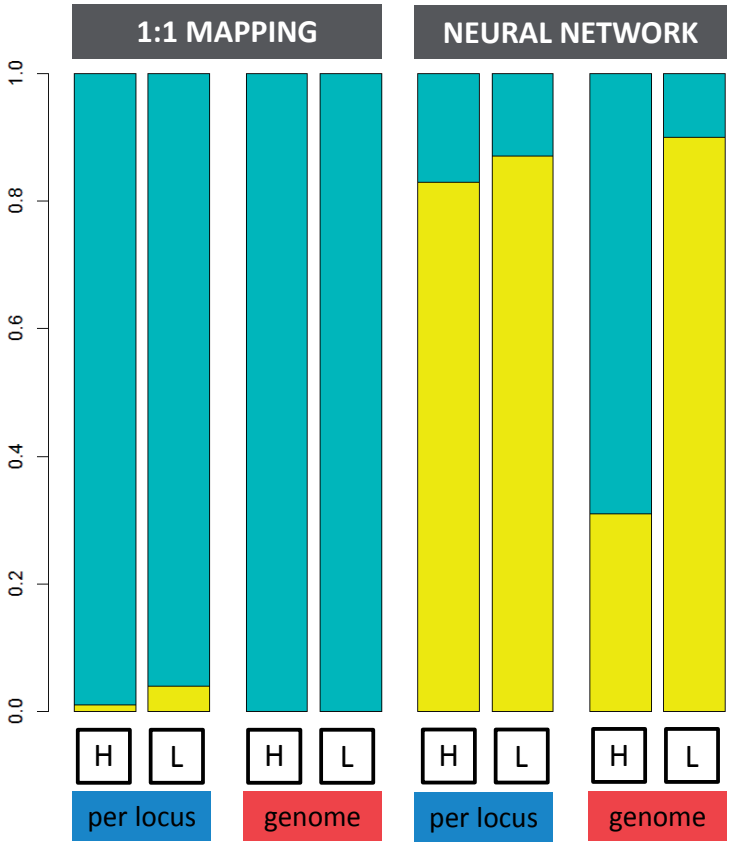

Figure 4.A6. Evolutionary outcome of the simulations of the ISD with alternative payoff configuration, for the different mechanistic implementations and mutation rates. The left bars of each pair show high mutation rates (H, $\mu = 0.001$), bars on the right show low mutation rates (L, $\mu = 0.0001$). Bars indicate the fraction of 100 replicate simulations ending up in each of the three ESSs of the ISD (Table 4.A2) after 100,000 generations.

to cycle is somewhat surprising, since the altered IPD has a cyclic attractor (Figure 4.A4) while the original game only had two pure-strategy attractors (Table 4.A1).

As in the original model, behavioural architecture and mutation regime strongly affect the outcome and dynamics of evolution. For instance, in the 1:1 mapping, the *Pavlov* equilibrium is a much more common outcome than in the neural network implementation. The *Pavlov* equilibrium was never observed to be invaded by any other strategy in any of the replicate simulations across all implementations. This suggests that given enough time, all simulations should end up in this equilibrium, the waiting time until this happens being determined by the mechanistic implementation. Consistently with this, the *Pavlov* equilibrium was more frequently observed for higher mutation rates. As for the original payoff configuration, the incidence of highly dynamic periods is lower for reduced mutation rates (especially in the neural network implementation).

Some of the simulation dynamics bear some resemblance to the cyclical attractor that was identified for this payoff configuration (periods with many subsequent peaks of *TFT*).

Figure 4.A6 shows the equilibria that were attained in the simulations of the ISD with alternative payoff configuration, again for all mechanistic implementations. This altered version of the ISD has two pure-strategy ESSs: *con-D* and *Pavlov* (similar to ESSs 1 and 2 in the original game). From the figure, it is obvious that the mechanistic implementation has a strong effect on the simulation outcome; in the 1:1 mapping, ESS 2 was by far the most common outcome, whereas in the neural network, ESS 1 was more common. The mutation regime also affected the outcome, especially in the neural network implementation, where ESS 2 was more common for entire-genome mutation than for per-locus mutation. As in the original game, ESS 2 was observed to succeed ESS 1, but never *vice versa*. Accordingly, one would expect to find ESS 2 more frequently in case of a higher mutation rate (since a higher mutation rate should lead to a faster transition from ESS 1 to ESS 2). Indeed, ESS 2 was observed less frequently for the lower mutation rate across all implementations. However, this effect is not very pronounced except for the neural network implementation with entire-genome mutation.